A video editing device includes a storage unit, a detection unit, and a composite unit. The storage unit stores a detection target, and a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another. The detection unit determines in accordance with the similarity determination condition whether or not the detection target is included in the original video content. The composite unit adds the message to the original video content in a case where a determination result by the detection unit indicates that the detection target is included in the original video content.
### Detection target table

<table>
<thead>
<tr>
<th>Search No.</th>
<th>File name</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. 1</td>
<td>aaa</td>
<td>construction</td>
</tr>
<tr>
<td>No. 2</td>
<td>aaa</td>
<td>construction</td>
</tr>
<tr>
<td>No. 3</td>
<td>ccc</td>
<td>construction</td>
</tr>
<tr>
<td>No. 4</td>
<td>ddd</td>
<td>person</td>
</tr>
<tr>
<td>No. 5</td>
<td>eee</td>
<td>person</td>
</tr>
<tr>
<td>m</td>
<td>mmm</td>
<td>object</td>
</tr>
</tbody>
</table>

**FIG. 2**
<table>
<thead>
<tr>
<th>Insertion position</th>
<th>Detected all frames</th>
<th>Detected all frames</th>
<th>Detected all frames</th>
<th>Detected all frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>This place seems to be XX gate.</td>
<td>What is shown is XX gate, and to the left and right of the XX gate...</td>
<td>Attractive constructions and landscapes remain at Asakusa.</td>
<td>A local character may be hidden in this scene.</td>
<td>The character which few persons know may be hidden into this scene.</td>
</tr>
</tbody>
</table>

![52 Condition table](image)

**FIG. 3**
<table>
<thead>
<tr>
<th>Frame No.</th>
<th>Search No.1</th>
<th>Search No.2</th>
<th>Search No.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>65</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>50</td>
<td>65</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n+1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Messages are set.

Similarity conditions are satisfied.

Number of detected frames condition is satisfied.
S101 Receive and store content to be edited

S102 Extract frame from content

S103 Cutting-out region from frame

S104 Extract feature from region

S105 Calculate similarity score from feature amounts of content and detection target

S106 Determine similarity from similarity determination condition

S107 Add message to content in insertion position

S108 End of process?

End

FIG. 5
VIDEO EDITING DEVICE, VIDEO EDITING METHOD, PROGRAM, AND MEDIUM IN WHICH THE PROGRAM IS RECORDED

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application is based upon and claims the benefit of priority from the prior Japanese Patent Application No. 2011-098226, filed on Apr. 26, 2011, the entire contents of which are incorporated herein by reference.

TECHNICAL FIELD

[0002] Embodiments of the invention relate to a video editing device for editing videos, a video editing method, a program and a medium in which the program is recorded.

BACKGROUND ART

[0003] In tourist spots such as theme parks and sightseeing areas, tourists shoot videos by using camcorders carried by themselves, and then record the shot videos as souvenir videos.

[0004] Editing of the videos of this kind by adding thereto information related to the videos can enhance the value of the original videos or make the videos impressive to the viewers. However, it is not easy for the people who shot the videos to edit the videos by themselves.

[0005] Meanwhile, in such tourist spots, videos of families or friends, or videos of characteristic landscapes, constructions or the like in the spots are often shot.

[0006] In this respect, a system to combine a shot video with a frame or a character of a theme park has been proposed (refer to JP, P2002-232832A).

[0007] Such conventional services give value to shot videos as souvenirs. The conventional services, however, can only combine an original video with a previously prepared image in a fixed manner, and thus cannot make the resultant video more valuable than a souvenir video.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1 shows a schematic configuration of a video editing system of a first embodiment;

[0009] FIG. 2 is a diagram showing a detection target table;

[0010] FIG. 3 is a diagram showing a condition table;

[0011] FIG. 4 shows an example of how a candidate for a message to be inserted on the basis of a detection result is set in a content; and

[0012] FIG. 5 is a flowchart showing operation of the video editing system.

DETAILED DESCRIPTION

[0013] According to one embodiment, a video editing device includes:

[0014] a storage unit to store a detection target, a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another;

[0015] a detection unit to determine in accordance with the similarity determination condition whether or not the detection target is included in the original video content; and

[0016] a composite unit to add the message associated with the detection target to the original video content in a case where a determination result by the detection unit indicates that the detection target is included in the original video content.

[0017] An embodiment will be described below in detail with reference to the drawings.

[0018] FIG. 1 is a diagram showing a configuration of a video editing system of the embodiment.

[0019] As shown in FIG. 1, the video editing system of the embodiment includes a video editing device 3, and an input unit 1 for setting editing conditions to the video editing device 3. The video editing device 3 reads an original video content (hereinafter referred to as an “original content”) including data such as video and sound from a recording medium 2, edits the original video content, and produces the edited video content (hereinafter, referred to as an “edited content”).

The recording medium 2 is an SD memory card, a Memory Stick or a DVD-ROM, for example. Additionally, a storage device such as an optical drive device, and a display device for displaying an editing target video are connected to the video editing device 3 when needed. Note that, “SD” and “Memory Stick” are registered trademarks.

[0020] The video editing device 3 previously records editing conditions for the original content and adds information such as a message to the original content in accordance with the editing conditions.

[0021] Examples of the editing conditions include a detection target, a detection condition, information to be added, and an insertion position of information. The input unit 1 is used to input such editing conditions to the video editing device 3.

[0022] The input unit 1 includes a keyboard used by a user to input a detection condition or the like to the video editing device 3, a network such as a LAN through which an image or sound of a person, a construction or the like stored in an external storage device is inputted to the video editing device 3 as a detection target, and a communication interface of the network.

[0023] The video editing device 3 includes a storage unit 30, a target registration unit 31, a condition setting unit 32, an additional information registration unit 33, a detection unit 34, a video reception unit 35, feature amount extraction units 36, 41, a similarity score calculation unit 37, a similarity determination unit 38, a frame extraction unit 39, a region cutting-out unit 40, a composite unit 42, and a video output unit 43.

[0024] The video editing device 3 is realized by a computer in which a memory, a CPU, a hard disk drive and a communication interface are installed, for example. The storage unit 30 is realized by the memory, the hard disk drive, or the like. A program installed in the hard disk drive, and the CPU, which reads this program into the memory and then executes the program, realizes the target registration unit 31, the condition setting unit 32, the additional information registration unit 33, the video reception unit 35, the feature amount extraction units 36, 41, the similarity score calculation unit 37, the similarity determination unit 38, the frame extraction unit 39, the region cutting-out unit 40, the composite unit 42, and the video output unit 43.

[0025] The storage unit 30 previously stores a detection target 50, a condition table 52, a message 54, and an original content 56.

[0026] The target registration unit 31 receives an image of a person or a construction that is desired to be detected, i.e., a detection target, and an attribute of the image, and stores the image and the attribute of the image in the storage unit 30 as the detection target 50. In addition, the target registration unit 31 records the file name and attribute of the image in the
detection target table \(50a\) (refer to FIG. 2) for each search number, which corresponds to the search order. Note that, in a case where the same detection target is searched for under different search conditions, the same search target is set in different search numbers among search numbers \(1 \text{ to } m\) in the detection target table \(50a\). The detection target \(50b\) includes a search target (image) and the detection target table \(50a\). The target registration unit \(31\) has a function to specify a region set to be a target in the search target.

[0027] The target registration unit \(31\) acquires data from a device connected via a LAN or a USB interface or acquires data from a recording medium by reading the data with use of a reading device such as a memory card reader or a DVD drive device. Note that, means for acquiring data from outside is not limited to the ones described herein, and the invention is not limited to the ones described in the embodiment.

[0028] As shown in FIG. 2, the file name and category (classification, type) of a detection target are associated with each other for each search number in the detection target table \(50a\).

[0029] In a search number 1, for example, the file name is “aaa” and the category is “construction.” In a search number 2, the file name is “aaa” and the category is “construction.” Moreover, in a search number 3, the file name is “ccc” and the category is “construction.” Furthermore, in a search number 4, the file name is “ddi” and the category is “person.” In the detection target table \(50a\), the search number 1 and the search number 2 correspond to the same detection target, but the detection conditions of the search number 1 and the search number 2 are different as will be described later.

[0030] Here, although the detection target is assumed to be an image, a video or sound may be set as the detection target by adjusting the feature amount extraction unit \(36\) to be described later. A plurality of images may be provided for a single object. For the purpose of further improving the detection accuracy, it is possible to use a plurality of images obtained by shooting a single object from various angles or a plurality of images obtained by shooting a single object with different brightness levels.

[0031] The condition setting unit \(32\) receives a condition for adding information to the original content, then records the condition in the condition table \(52\) and stores the condition table \(52\) in the storage unit \(30\). To put it specifically, the condition setting unit \(32\) can edit a similarity determination condition in accordance with a request.

[0032] As shown in FIG. 3, in the condition table \(52\), a similarity determination condition (a similarity score and the number of detected frames), information to be added (hereinafter, described as a message), and an insertion position in association with one another are stored for each search number. The similarity determination condition is used as the condition for adding the message. The insertion position shows a frame of the video to which the message is added.

[0033] To put it specifically, the condition table \(52\) stores, for each search number, the similarity determination condition for determining the similarity of the original content to the detection target, the message to be added to the original content when this condition is matched, and the insertion position for adding the message to the original content, in association with one another. In the condition table \(52\) shown in FIG. 3, the search number 1 and the search number 2 correspond to the same detection target, but different similarity determination conditions and different messages are set in the search numbers 1 and 2, respectively. Accordingly, a different message can be added to the original content in accordance with the similarity to the detection target.

[0034] The similarity determination condition is inputted to the video editing device \(3\) via a keyboard connected to the video editing device \(3\). Meanwhile, it is also possible to display a graphic user interface (GUI) on the screen of a computer (input terminal) connected to the video editing device \(3\) via a LAN or the like, and to allow input of the condition to the video editing device \(3\) via the GUI, for example.

[0035] The additional information registration unit \(33\) receives the message to be added to the original content as well as the insert position where the message is inserted and records the message as well as the insert position in the condition table \(52\) and the message \(54\). The condition table \(52\) and the message \(54\) are recorded in the storage unit \(30\). The message is, for example, a sentence to be displayed as a caption or a sentence to be superimposed on the video.

[0036] The additional information registration unit \(33\) acquires data from a device connected via a LAN or a USB interface, or reads data from a medium such as an SD memory card or a CD-ROM, or receives data from a computer connected via a LAN or the like. The method for inputting data and the method for receiving data are not limited to the ones described in the embodiment.

[0037] The video reception unit \(35\) receives an original content from outside of the video editing device \(3\) and saves the original content in the storage unit \(30\). The video reception unit \(35\) acquires an original content from a device connected via a LAN or a USB interface, or reads an original content from the recording medium \(2\) such as an SD memory card or a DVD. The method for receiving an original content is not limited to the ones described in the embodiment.

[0038] Various video formats such as MPEG-2, 3GP, AVI and H.264/MPEG-4 AVC are assumed to be used as the formats of original contents. Note that, the formats are not limited to the ones mentioned herein. An original content includes an image frame.

[0039] Moreover, it is possible to perform primary editing on a video shot by a camcorder and then to input the edited video to the video editing device \(3\) as an original content. Moreover, the video editing device \(3\) may include a primary editing function and produce a digest video by deleting an unnecessary portion from the original content. Here, the video editing device \(3\) performs secondary editing on the digest video and adds a message to the digest video.

[0040] The feature amount extraction unit \(36\) calculates (extracts) a feature amount \(51\) of a detection target from the detection target \(50\) stored in the storage unit \(30\) and stores the extracted feature amount in the storage unit \(30\).

[0041] The method for extracting a feature amount is not limited to any particular method. The method for extracting a feature amount of a recognition target from a color image, for example, may be color-CoHOG (Co-occurrence Histograms of Oriented Gradients), CoHED (Co-occurrence Histograms of Pairs of Edge Orientations and Color Differences), CoH (Co-occurrence Histograms of and Color Differences), and the like. However, the method for extracting a feature amount in the invention is not limited to these methods.

[0042] The detection unit \(34\) includes the similarity score calculation unit \(37\) and the similarity determination unit \(38\). The detection unit \(34\) determines, in accordance with the similarity determination condition set for each search number in the condition table \(52\), whether or not a detection target is included in the original content \(56\) and notifies the composite unit \(42\) of the determination result.
[0043] The similarity score calculation unit 37 compares a feature amount 55 of a frame of the original content extracted by the feature amount extraction unit 41 and the feature amount 51 of the detection target and thereby calculates a similarity score indicating a degree of similarity therebetween.

[0044] The similarity determination unit 38 determines the similarity between the images in accordance with the similarity score calculated by the similarity score calculation unit 37 and the similarity determination condition read from the condition table 52.

[0045] The similarity determination unit 38 compares the calculated similarity score with the similarity score of the determination condition. When the calculated similarity score satisfies the similarity score of the determination condition, the similarity determination unit 38 further evaluates the frame in terms of the condition indicating a required number of detected frames.

[0046] The similarity determination unit 38 performs the similarity determination by checking the calculated similarity score against the similarity determination condition in the condition table 52. In a case where the original content matches the determination condition, the similarity determination unit 38 outputs the result of similarity determination, the message to be added, and the insertion position of the message (time axis information of the video) to the composite unit 42. Note that, the similarity determination unit 38 may notify the composite unit 42 of the search number and the result of similarity determination, and the composite unit 42 may read the message and the insertion position from the message 54.

[0047] For the search number 1, for example, upon detection of five consecutive frames each having a similarity score S of 65%, the similarity determination condition for the search number 1 is satisfied. This similarity determination condition requires “a moderate similarity score and also a moderate number of detected frames.” In this case, the similarity determination unit 38 outputs the result of similarity determination, a probable message, i.e., “this place seems to be XX gate,” and the insertion position.

[0048] In addition, for the search number 2, upon detection of at least eight consecutive frames each having a similarity score S of 85% or higher, the similarity determination condition for the search number 2 is satisfied. This similarity determination condition requires “a high similarity score and a large number of detected frames.” In this case, the similarity determination unit 38 outputs the result of similarity determination, a definite message, i.e., “what is shown is XX gate, and to the left and right of XX gate . . . ,” and the insertion position.

[0049] Furthermore, for the search number 4, upon detection of at least ten consecutive frames each having a similarity score S exceeding 30%, the similarity determination condition for the search number 4 is satisfied. This similarity determination condition requires “a low similarity score but a large number of detected frames.” In this case, the similarity determination unit 38 outputs the result of similarity determination, a message prompting checking, i.e., “a local character may be hidden in this scene,” and the insertion position.

[0050] A message that arouses the interest of the viewers is registered in the condition table 52 in accordance with the detection target and the similarity determination condition.

[0051] In a case where there are a plurality of candidate messages, it is conceivable to output a message corresponding to a similarity determination condition that requires a higher similarity, or to simultaneously output the plurality of messages. In addition, the insertion position of a message is previously set in the case of the above description. However, it is also possible to add a message to all of the frames satisfying the similarity score condition in the determination condition without providing the insertion position, for example.

[0052] Moreover, the techniques used for the feature amount extraction method and for the similarity determination are known by JP, P2011-18238A, JP, P2009-119460A, for example. These techniques can be used in the embodiment. Note that, the techniques shown herein are only examples, and there are a feature amount extraction technique and a similarity determination technique other than the techniques mentioned herein, and any method can be used.

[0053] The frame extraction unit 39 extracts one by one, the frames of an original content stored in the original content 56 in the storage unit 30. The type of frame serving as the key is MPEG-2 I frame or the like, for example, and other than the I frame, a frame serving as a representative for each scene may be extracted, for example.

[0054] The region cutting-out unit 40 cuts out a region for comparison with the feature amount 51 of the detection target previously registered in the storage unit 30 from the frame extracted by the frame extraction unit 39.

[0055] The feature amount extraction unit 41 extracts a feature amount from the image cut out by the region cutting-out unit 40 and stores the feature amount in the feature amount 55 in the storage unit 30.

[0056] The composite unit 42 adds the message to the original content on the basis of the determination result indicating that the similarity determination condition is satisfied, the message, and the insertion position, which are outputted from the detection unit 34 (similarity determination unit 38). The composite unit 42 produces an edited content by adding the message to the original content as a caption and stores the edited content in the edited content 57 in the storage unit 30.

[0057] To put it specifically, on the basis of the determination result indicating that the detection target is included in the original content, i.e., that the original content satisfies the similarity determination condition, the composite unit 42 reads the message corresponding to the similarity determination condition from the storage unit 30 and adds the message to the original content. To put it more specifically, the composite unit 42 takes into consideration the number of detected frames each having a similarity score to the detection target which is equal to or greater than a predetermined value, or the consecutiveness of the frames, and then adds the message to the original content.

[0058] For adding a message to an original content, a method to combine a message with a content by superimposing the message on the content is available, for example. Moreover, in a case where a caption area is provided as in the case of DVD, Blu-ray, or the like, a method to set (insert) the message in the area is available. In a case where a message is set in the caption area, operation to turn ON/OFF the displaying of the message is available during viewing of the content. Any of the methods can be used for adding a message. Blu-ray is a registered trademark. Meanwhile, addition of a message to an original content does not have to be necessarily per-
formed by adding the message as a caption and may be performed by superimposing the message onto the screen of the video.

**[0059]** The video output unit 43 sends the edited content combined with the message to the outside of the video editing device 3. The edited content can be sent via a network connection using a LAN, or via a recording medium such as a CD-ROM, a DVD-ROM, or a Blu-ray disc, for example.

**[0060]** Operation of the video editing system of the embodiment will be described below with reference to FIG. 4 and FIG. 5. The video editing system of the embodiment prepares for editing an original content, first.

**[0061]** Upon input of an image of a person, a construction and the like that are desired to be detected to the video editing device 3 as detection targets by the input unit 1, the target registration unit 31 receives and then stores the detection target in the storage unit 30. In addition, the target registration unit 31 writes the detection target and the attribute of the detection target in the detection target table 50a for each search number in accordance with the input order of the detection targets.

**[0062]** In addition, upon input of a condition for adding a message to the video editing device 3 when a detection target is inputted to the video editing device 3, the condition setting unit 32 sets the condition in the condition table 52 as a similarity determination condition.

**[0063]** For example, the condition table 52 is set to have, as the similarity determination condition, a similarity score condition in terms of a similarity score to be calculated from the feature amount of an original content and the feature amount of a detection target, and a number-of-detected-frames condition in terms of the number of frames satisfying the similarity score condition. The number-of-detected-frames condition indicates a required number of consecutive frames satisfying the similarity score condition.

**[0064]** Moreover, upon input of a message to be added to an original content and an insertion position to the video editing device 3 by the input unit 1, the addition information registration unit 33 receives the message and the insertion position and sets the message and the insertion position in the condition table 52 and also stores the message and the insertion position in the message 54 of the storage unit 30.

**[0065]** The feature amount extraction unit 36 extracts a feature amount from a detection target stored in the storage unit 30 and stores the feature amount in the original content in the storage unit 30. Accordingly, the preparation for editing the original content ends.

**[0066]** Next, editing processing of an original content will be described.

**[0067]** Upon input of an original content to the video editing device 3 from outside, the video reception unit 35 receives the original content and stores the received original content in the original content 56 in the storage unit 30 (step S101 in FIG. 5).

**[0068]** The feature extraction unit 39 extracts one frame from the original content stored in the storage unit 30 (step S102). The type of frame serving as the key is an MPEG-2 I frame, for example. Meanwhile, a frame serving as a representative for each scene may be extracted, for example.

**[0069]** The region cutting-out unit 40 cuts out a region for comparison with a previously registered feature amount of a detection target from the frame extracted by the frame extraction unit 39 (step S103).

**[0070]** The feature amount extraction unit 41 extracts a feature amount from the image cut out by the region cutting-out unit 40 and stores the feature amount in the feature amount 55 in the storage unit 30 (step S104).

**[0071]** Subsequently, the similarity score calculation unit 37 reads the feature amount of the original content and the feature amount of the detection target from the storage unit 30, and then compares the two feature amounts and thereby calculates the similarity score therebetween on the basis of the degree of matching between the feature amounts (step S105).

**[0072]** The similarity determination unit 38 determines the similarity between the images on the basis of the similarity score calculated by the similarity score calculation unit 37 and the similarity determination condition in the condition table 52 (step S106).

**[0073]** How the similarity is determined in step S106 will be described with reference to FIG. 4. FIG. 4 shows similarity of the current frame and the previous frames for each search number. Here, a case of determining whether or not the frame with the frame number 5 matches the similarity determination condition of the search number 1 will be described.

**[0074]** According to the condition table 52, the similarity score condition is that the similarity score S is 85%≥S≥60% and also the number-of-detected-frames condition is N≥5 in the similarity determination condition of the search number 1. The frame with the frame number 5 has 65% of similarity score to the detection target of the search number 1 and thus satisfies the similarity score condition. Moreover, any of the nearest five frames including the frame with the frame number 5 satisfies the similarity score condition. Accordingly, the frame with the frame number 5 satisfies the number-of-detected-frames condition. Since the frame with the frame number 5 satisfies the similarity score condition and the number-of-detected-frames condition, the similarity determination unit 38 determines that the frame with the frame number 5 matches the similarity determination condition of the search number 1. Thus, the message of the search number 1 is set as an insertion message candidate as shown in FIG. 4.

**[0075]** Note that, in a case where there are a plurality of search numbers, the processing from step S103 to step S106 is performed for each of the search numbers. Accordingly, as shown in FIG. 4, the similarity to the detection target set in each of the search numbers 1 to m is calculated for the frame with each frame number.

**[0076]** Upon end of the determination of the similarity between the images of the frames and the detection targets of the respective search numbers, the detection unit 34 notifies the composite unit 42 of the determination result. To put it specifically, in a case where a frame matches the similarity determination condition, the detection unit 34 outputs the result of similarity determination, the message to be added and the insertion position to the composite unit 42.

**[0077]** As an upper level determination, the detection unit 34 detects whether or not a detection target is included in the original content in accordance with a similarity determination condition in the condition table 52.

**[0078]** The composite unit 42 adds the message to the original content in accordance with the determination result indicating that the detection target is included in the original content, the message and the insertion position (step S107). Alternatively, upon notification of the search number and the detection result, the composite unit 42 adds the message to the
original content in accordance with the message and the insertion position which are read from the message 54 in the storage unit 30 (step S107).

[0079] Subsequently, when there is no instruction to end the processing (No in step S108), the processing from step S103 to step S107 is performed for the next frame. Upon end of the processing for all of the frames (Yes in step S108), the video output unit 43 writes the edited content to which the message has been added (combined) to a recording medium (a CD-R, a DVD-R, each type of a memory card or the like) for external output. As the recording medium, a CD-R, a DVD-R, each type of a memory card or the like is used. Note that, in a case where editing is performed on the basis of a request received via a network or the like, the edited content is sent to the originator of the editing request. In addition, addition of the information (message) to the original content does not have to be necessarily performed as addition of a caption, and the information may be superimposed onto the screen of the video.

[0080] Here, a service utilizing the video editing system of the embodiment will be described. According to the video editing system of the embodiment, a previously inputted detection target and images included in an original content are compared. Then, a message is added to a frame of a similar image, and the edited content is outputted. Thus, the video editing system of the embodiment can provide the following service.

[0081] An example of a case where the video editing system of the embodiment is applied to a video shot in a theme park will be described. The video editing system of the embodiment is applied to an original video content in which a family or a friend is shot together with a character of the theme park. In a case where it is detected that, in the original video content, frames each having a similarity score to an image of the character that exceeds a certain similarity score are consecutive for at least a certain period of time or at least a certain number of frames, the video editing system adds a message just like the contents which the character of the theme park speaks to the video to the frames in which the character is detected. Thus, a new sensation can be added to the original video content.

[0082] In addition, the video editing system of the embodiment can add a message to a video entered in a video contest with a predetermined theme. In a case where a target image designated by the host of the contest exists in a video entered in the contest, for example, the video editing system outputs a video in which a winning notification message is added to the frame in which the target image is shown.

[0083] Moreover, in a case where at least a certain number of frames each having a high similarity score to a representative building or construction in a tourist spot are detected in an original video content shot in the tourist spot, the video editing system of the embodiment adds a message such as an explanation for the building or construction to the frames.

[0084] Moreover, it is also possible to register a video of a predetermined product to the video editing system of the embodiment in advance so that the video editing system of the embodiment can add information about the product when a video of the product is detected in an original video content.

[0085] Furthermore, when a video of a registered person is detected in a video shot by a monitoring camera installed in a monitoring target area where a particular person enters and leaves, the video editing system of the embodiment adds the attribute and name of the detected person to the output video and displays the video.

[0086] As described above, according to the video editing system of the embodiment, new value can be added to an original video. In addition, it is possible to almost automatically execute a series of processing operations including reception of an original content, creation of an edited content by adding a message to the original content, and delivery (sending) of the created edited content to the originator of the request.

[0087] As described above, according to the above described embodiment and the application of the embodiment, there are provided a video editing device, a video editing method, a program, and a recording medium in which the program is recorded, which are capable of adding new value to a shot video.

[0088] While certain embodiments have been described, these embodiments have been presented by way of example only, and are not intended to limit the scope of the inventions. Indeed, the novel embodiments described herein may be embodied in a variety of other forms; furthermore, various omissions, substitutions and changes in the form of the embodiments described herein may be made without departing from the spirit of the inventions. The accompanying claims and their equivalents are intended to cover such forms or modifications as would fall within the scope and spirit of the inventions.

[0089] Furthermore, each configuration element described in the above embodiment may be implemented by a program installed in a storage such as a hard disk drive of a computer, or the functions of the present invention may be implemented by a computer in such a way that the above program is stored in a computer-readable electronic medium and then read by the computer from the electric medium.

[0090] Examples of the recording medium include recording media such as a CD-ROM and a DVD-ROM, as well as a flash memory, removable media and the like. Moreover, the configuration elements may be distributed and stored in different computers connected to each other via a network and may be implemented by communications between the computers functioning the configuration elements.

What is claimed is:
1. A video editing device comprising:
   a storage unit to store a detection target, a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another;
   a detection unit to determine in accordance with the similarity determination condition whether or not the detection target is included in the original video content; and
   an output unit to output the detection result in the case where a determination result by the detection unit indicates that the detection target is included in the original video content.
2. The video editing device according to claim 1, wherein the storage unit further stores an insertion position where the message is inserted in the original video content, and the composite unit inserts the message in the insertion position of the original video content on the basis of the insertion position.
3. The video editing device according to claim 1, wherein the original video content includes image frames, the similarity determination condition includes a similarity score condition and a number-of-frames condition indicating a required number of frames satisfying the similarity score condition, and the detection unit determines whether or not the detection target is included in the original video content in accordance with the similarity score condition and the number-of-frames condition.

4. The video editing device according to claim 1, wherein the detection unit compares a feature amount of the original video content and a feature amount of the detection target, thereby calculates a similarity score and determines a similarity therewith between by checking the similarity score against a previously registered similarity determination condition.

5. The video editing device according to claim 3, wherein in a case where frames satisfying the similarity score condition are consecutive in number that satisfies the number-of-frames condition, the composite unit adds the message to the original video content.

6. The video editing device according to claim 5, wherein in a case where frames satisfying the similarity score condition are consecutive in number that satisfies the number-of-frames condition, the composite unit adds the message to the frames of the original video content that satisfy the similarity score condition.

7. A video editing device comprising:
   a storage unit to store a detection target, a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another;
   a detection unit to determine in accordance with the similarity determination condition whether or not the detection target is included in the original video content;
   an composite unit to add the message associated with the detection target to the original video content in a case where a determination result by the detection unit indicates that the detection target is included in the original video content;
   a target registration unit to receive the detection target and to store the received detection target in the storage unit;
   a condition setting unit to receive the similarity determination condition and to store the received similarity determination condition in the storage unit;
   an additional information registration unit to receive the message and to store the received message in the storage unit;
   a video reception unit to receive the original video content and to store the received original video content in the storage unit; and
   a video output unit to output the original video content to which the message is added.

8. A video editing method comprising:
   storing a detection target, a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another;
   determining in accordance with the similarity determination condition whether or not the detection target is included in the original video content; and
   adding the message to the original video content in a case where it is determined that the detection target is included in the original video content.

9. A program causing a computer to execute processing, the program causing the computer to function as:
   a storage unit to store a detection target, a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another;
   a detection unit to determine in accordance with the similarity determination condition whether or not the detection target is included in the original video content; and
   an composite unit to add the message to the original video content in a case where a determination result by the detection unit indicates that the detection target is included in the original video content.

10. A medium in which a computer-readable program is recorded, the program causing the computer to function as:
    a storage unit to store a detection target, a similarity determination condition for determining a similarity of an original video content to the detection target, and a message in association with one another;
    a detection unit to determine in accordance with the similarity determination condition whether or not the detection target is included in the original video content; and
    an composite unit to add the message associated with the detection target to the original video content in a case where a determination result by the detection unit indicates that the detection target is included in the original video content.

* * * * *